MAXIMUM TEMPERATURE PREDICTION

by R. Maine

Divisional Office, Bureau of Meteorology, Adelaide

(Manuscript received June, 1958)

Abstract: A set of 74 January daily maximum temperatures at Adelaide, South Australia, can be shown to correspond with reasonable precision to an empirically predicted set based on three homogeneous governing parameters available at 6 a.m. local time. Selection of governing parameters is accomplished statistically.

1. INTRODUCTION

This report has been produced because of the need to establish an objective procedure with which to forecast daily maximum temperatures, using in the first instance homogeneous early morning data.

As maximum temperatures are essential in the preparation of fire weather warnings, direct application can be immediately found in this field, and as most Bureaux are responsible for the issue of such warnings to Government Forestry reserves the general approach can provide valuable objective assessment of the maximum temperature factor for a comparatively economical initial outlay in time. In fact, the procedure has been applied to Mt. Gambier in the southeast of South Australia where weather conditions closely approximate, or are simply and systematically different from, those in two nearby forestry reserves, namely Mt. Burr and Penola. In this case only two parameters were governing, but the investigation was not exhaustive and errors may be further reduced by incorporation of further though less significant variables. However, the small amount of data used was a valuable guide to forecasting days with extreme temperature range. In the test case Adelaide data which were readily accessible were used, as handling of a large number of parameters is usually inevitable in the early stages to determine relative significance. Also publication of capital maximum temperature forecasts in local newspapers makes the use of this data more important.
2. REVIEW OF LITERATURE

English references to maximum temperature forecasting do not appear very extensive and many of the methods proposed are not suitable for local application.

Probably by far the best known and quickest approach to maximum temperature estimation is Gold's (1933) which basically sets an upper limit to the value of the maximum in relation to the lapse rates in the lowest six or seven thousand feet of the atmosphere. The application of this method to the present problem is made difficult by the absence of any night time radiosonde ascent locally, and by the fact that the super-adiabatic lapse rates frequently occur in the lowest layers. A more elegant computation is made quantitatively by assessing total heating from available radiation and advection and applying this to the night time sounding.

Neiburger (1941) computes a value for the maximum on similar lines to Gold though making allowance for a greater number of variables affecting total heat input.

Brunt (1932) discussed the form of the diurnal surface temperature curve at the ground but again when operating with the resulting mathematical expression many obstacles are encountered which require excessive simplification before extracting a practical answer. In particular local soil constants are practically unknown, and considerable added variation is possible with approximate values for different areas; (Johnson and Davies, 1927). It is understood however that the Waite Agricultural Institute is undertaking a study of soil constants in parts of the Adelaide Plains.

Under the circumstances it appears that a deterministic approach although eventually essential should not be preferred, for the present, to an empirical method producing useable results, and indeed such a method provides useful information in any precise approach to the problem.

As it is fairly clear that more than one variable will be involved in relation to maximum temperatures natural recourse is immediately taken to the established practices of multiple regression or correlation.

Unfortunately, multiple linear regression techniques are on the whole unwieldy with large samples and even more so when any departure from linearity or scalar quantities is involved, and are therefore not well adapted to this type of investigation. However, this technique was used a number of times to test assumptions, but using only small samples. Many methods of multiple correlation are
available and probably the most suitable is that of graphical correlation as the functional relations between variables need not be known or assumed. Several varieties of graphical procedures exist and a final choice was made between the method of deviations, co-axial plotting (see Linsley, Kohler and Paulhus, 1949), and graphical integration (Bundgaard, 1951). The method of deviations was ultimately used as it is simple to apply and is uncomplicated by introduction of vector quantities or variables of changing sign. In addition a second approximation can readily be made if required.

The basic underlying assumption with the application of the results of the correlation analysis to forecasting is of course that identical weather conditions, in respect to the determining variables will produce the same dependant variable which may be for example temperature range in all cases.

3. TREATMENT OF DATA

The first step after examining the available methods of attack was to attempt to isolate those variables having the strongest relation to the maximum temperature and to assess their degree of correlation, i.e. put them into some order of importance. From a physical standpoint and employing ordinary "synoptic" experience many factors can be suggested but their effect in relation to one another is uncertain. Consequently a time series of maximum temperature, 6 a.m. temperatures, mean cloud from 6 a.m. to 3 p.m., 6 a.m. M.S.L. pressure, 6 a.m. dew-point, surface wind, and other variables were plotted for a section of the summer of 1955.

It was immediately evident that a high degree of correlation existed between 6 a.m. temperatures and maxima. The goodness of fit was such that when plotted in a scatter diagram the relation

\[ T \text{ (MAX)} = T \text{ (6 a.m.)} + K \tag{1} \]

for \( K = 24 \) and \( T \) = temperature in °F

was a reasonable estimate for midsummer clear days.

Other variables when plotted in scatter diagrams revealed some connection with cloud and wind. Accordingly it was decided that if 6 a.m. data were available, or data about that time, the difference between the 6 a.m. and the maximum temperature (hereinafter referred to as the range \( R \)) could be regarded as the variable requiring to be forecast. Under existing arrangements at this Bureau, early morning observations from most State forestry reserves are reported.
A working sample of about 50 observations was chosen from January and February, 1955, excluding those days on which precipitation was recorded between 6 a.m. and 6 p.m., and those days on which a major air mass change occurred between 6 a.m. and 6 p.m. (but not excluding sea breezes).

The range R was then plotted in scatter diagrams against various factors, for example, available insolation, mean cloud amount, 6 a.m. dew point, 1700Z 2000 ft winds, 6 a.m. temperature, etc.

Berry, Bollay and Beers (1945) gave a good account of the important variables in this respect; however, strengths of correlations were not clearly indicated.

Three variable relationships were also plotted with the strongest determining variates, but as they stood they were too weak to draw any definite conclusions about relative importance.

In all scatter diagrams the strength of relationships was noted visually by the degree of scatter. In most cases this appears quite satisfactory.

The main information derived from the scatter diagrams was that available insolation was the most strongly related variable, with wind and temperature as next strongest, respectively.

Available insolation was calculated as a percentage of clear sky insolation using the short wave cloud albedo figures of Fritz (1950). It appears from experience that the percentage reduction due to cirrus or cirrostratus is too high.

Using these calculated curves the relationship of insolation to range could be shown to fit linearly, in the mean, an expression

\[ R = 0.32H - 8 \]  

(2)

where H is available insolation, expressed as a percentage of clear sky insolation for the sample in question.

An attempt to introduce temperature range R as a predictor produced results which were not significant. It was then decided to group the variables which were observed to be related to R and insert them into a multiple linear regression. The magnitude of "t" obtained in the test of significance of co-efficients is used as an indicator to order of strength of relation, from a linear standpoint at least. Accordingly, with range R as the stochastic variate a multiple regression based on 30 observations (mid January to mid February, 1955) was computed using variates which when arranged in descending order of "t" took the following order:
1. Available insolation
2. Northerly wind component
3. 6 a.m. temperature
4. Easterly wind component
5. 6 a.m. mixing ratio
6. 24 hour change in 6 a.m. temperature

With the inclusion of several uninformative variates, e.g. 5 and 6 in the regression, it might be anticipated that the significance of any informative variate might be considerably reduced. Indeed this was the case and no value of "t" for any co-efficient was significant in the above regression.

A graphical correlation using the method of deviations was then carried out to the first approximation using only the first four variates, wind vector being plotted as second variate. To avoid the non-objectivity of mean cloud for the day in any forecast to be made at 6 a.m., the 6 a.m. cloud was inserted in its place - a step which did not significantly change the final result.

The resulting set of correlation diagrams based on 30 observations from mid January to mid February, 1955, was then used as a predictor of maxima during January, 1957.

The characteristics of the distribution of errors in the predicted set of maxima were the following:

Bias of errors $+0.3^\circ F$

Mean modulus of errors $2.4^\circ F$

Standard error $3.0^\circ F$

Range $-5^\circ F$ to $+4^\circ F$

As the method had so far yielded encouraging results it seemed worthwhile to revise the figures for available insolation, pool the data for 1955 and 1957 and replot diagrams to the second approximation.

Regarding the relation of insolation to cloud, observational results of incident transmitted energy seemed more realistic in these calculations, than the use of albedo figures published for different cloud types.
Many empirical relations between insolation and cloud exist, but the observational data of Haurwitz (1945) appeared to be most suitable, and these figures were used to evaluate a figure of percentage reduction of clear day midsummer insolation in relation to cloud amount and type.

This percentage reduction of the midsummer, noon, clear day beam, is mainly a function of cloud amount and type. Over the summer months (here taken as October to March) the variation of this reduction due to time of year is not more than a few percent, so any relation established for one month could be applied to data in other months equally as well.

Haurwitz's figures on the percent of average total daily insolation (sun plus sky) passed by various cloud amounts were then replotted for cloud densities. The Smithsonian Institution Table 152 (6th edition) was then used to interpolate cloud type on the cloud density scale for overcast conditions and the curves of incident insolation were then completed to the shape of the curves observed by Haurwitz. The result was the first of the 4 graphs of the prediction diagrams Fig. 1a.

It is interesting from the point of view of any further attempt at linear regression that the curves worked up from the above data are approximately curved logarithmically, i.e. straight lines when plotted on arithmetic, log. co-ordinates.

Having established a suitable basis for the percentage reduction of clear day insolation with cloud amount and type, the maximum temperature data for 1955 and 1956 were pooled and the correlation diagrams calculated to the second approximation.

The error distribution of range forecasts then had the following characteristics:

Bias $+0.1^\circ F$

Mean modulus $1.5^\circ F$

Standard error $2.4^\circ F$

Range $-9^\circ F$ to $+6^\circ F$
Fig 1a. Relation between cloud amount and per cent reduction of clear day average total insolation.

Fig 1b. Relation between $R'_{1}$ (first estimate of range) and per cent reduction of clear day average insolation.

Fig 1c. Relation between correction $D_{1}$ and 6 a.m. temperature.
Fig 1d. Curves of correction $D_2$ as related to 1700.Z. - 2000 ft wind.
The analysis of this data in relation to the variables, 6 a.m. temperature, 6 a.m. cloud, 1700Z wind vector, was capable of accounting for 83 per cent of the natural variance of the maximum temperature itself. Of the residual 17 per cent, two observations of the total number of 66 contributed 5 per cent. These two cases, one in error of +9°F and the other -6°F were examined, and in the first case a large positive error occurred when cloud rapidly developed from a clear sky at 6 a.m.; and in the other negative error occurred when overcast conditions at 6 a.m. broke to almost clear by 9 a.m. It therefore seems justifiable to surmise that a correct degree of subjective judgment in relation to cloud variation would result in a reduction of forecast error variance for cases where an abnormally fast clouding over or clearing is expected.

It appears that there will be some practical lower limit to residual variance, as at each inclusion of a new variable it becomes increasingly difficult to effect a given reduction in the residual variance.

From experience with fire weather warnings a forecast maximum temperature within the range of 2°C about the true value can be considered as a practical success, all other values meeting with varying degrees of failure. These limits may appear restrictive but it must be remembered that the other variables in a fire weather forecast, e.g. dew point and wind speed, also contribute their own error variances, and therefore allowance for these must be made when computing a composite rating such as the Luke fire danger index. Such an index, which can be used in relation to fire spread rate and wind direction is the real meteorological tool of a fire fighting organisation.

The pooled prediction graphs were subsequently used on the months January and February, 1958, as a test sample. The list of errors is reproduced in Table 1.

It can be seen that the method commences to break down during February - a fairly natural consequence as the base data are drawn predominantly from January. Consequently it was decided to revise all data and use January observations only.

There were 74 observations available in all and these were then used to recompute all correlation diagrams. All new diagrams took the same pattern but with certain systematic alterations on previous similar diagrams. In this set of data the correlation diagram of percentage reduction of insolation was plotted first, followed by 1700Z 2000 ft wind, and 6 a.m. temperature. The last named correlation being weak but nevertheless identifiable.
### Table 1. Errors in maximum temperature forecasts, January and February, 1958

<table>
<thead>
<tr>
<th>Date</th>
<th>Observed Maximum Temperature (°F)</th>
<th>Estimated Maximum Temperature (°F)</th>
<th>Range (°F)</th>
<th>Maximum Temperature (°F)</th>
<th>Date</th>
<th>Observed Maximum Temperature (°F)</th>
<th>Estimated Maximum Temperature (°F)</th>
<th>Range (°F)</th>
<th>Maximum Temperature (°F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>January 1</td>
<td>-2</td>
<td>13</td>
<td>71</td>
<td></td>
<td>January 21</td>
<td>-1</td>
<td>9</td>
<td>72</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>22</td>
<td>74</td>
<td></td>
<td>January 22</td>
<td>+3</td>
<td>12</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-3</td>
<td>23</td>
<td>81</td>
<td></td>
<td>January 23</td>
<td>-11</td>
<td>32</td>
<td>79</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+2</td>
<td>24</td>
<td>87</td>
<td></td>
<td>January 24</td>
<td>+4</td>
<td>26</td>
<td>81</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-1</td>
<td>25</td>
<td>92</td>
<td></td>
<td>January 25</td>
<td>+5</td>
<td>23</td>
<td>86</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+1</td>
<td>26</td>
<td>88</td>
<td></td>
<td>January 26</td>
<td>+2</td>
<td>29</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+1</td>
<td>27</td>
<td>89</td>
<td></td>
<td>January 27</td>
<td>-3</td>
<td>23</td>
<td>86</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+2</td>
<td>28</td>
<td>95</td>
<td></td>
<td>January 28</td>
<td>-15</td>
<td>30</td>
<td>84.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-5</td>
<td>29</td>
<td>102</td>
<td></td>
<td>January 29</td>
<td>-12</td>
<td>24</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+1</td>
<td>30</td>
<td>74</td>
<td></td>
<td>February 1</td>
<td>+9</td>
<td>18</td>
<td>78</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+1</td>
<td>31</td>
<td>75</td>
<td></td>
<td>February 2</td>
<td>-4</td>
<td>33</td>
<td>89</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-1</td>
<td>32</td>
<td>79</td>
<td></td>
<td>February 3</td>
<td>-3</td>
<td>18</td>
<td>95</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+2</td>
<td>33</td>
<td>95</td>
<td></td>
<td>February 4</td>
<td>-1</td>
<td>23</td>
<td>76</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-8</td>
<td>34</td>
<td>87</td>
<td></td>
<td>February 5</td>
<td>+6</td>
<td>30</td>
<td>77</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+3</td>
<td>35</td>
<td>95</td>
<td></td>
<td>February 6</td>
<td>+1</td>
<td>22</td>
<td>84</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-1</td>
<td>36</td>
<td>80</td>
<td></td>
<td>February 7</td>
<td>+1</td>
<td>21</td>
<td>91</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-8</td>
<td>37</td>
<td>87</td>
<td></td>
<td>February 8</td>
<td>-3</td>
<td>18</td>
<td>93</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+2</td>
<td>38</td>
<td>72</td>
<td></td>
<td>February 9</td>
<td>-3</td>
<td>20</td>
<td>98</td>
<td></td>
</tr>
<tr>
<td></td>
<td>+6</td>
<td>39</td>
<td>73</td>
<td></td>
<td>February 10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* January 9 change at 1330 CST
* January 14 change at 1200 CST
* January 18 change at 1200 CST
* January 23 change at 1200 CST

* January 28 change at 1200 CST
* January 29 rain

February 5 heavy smoke pall from Kangaroo Island fire.
The prediction diagrams are reproduced in Figs. 1b, c and d. Instructions for the use of Figs. 1a, b, c and d are as follows:

The percentage reduction of insolation is obtained by entering Fig. 1a given 6 a.m. cloud amount and type. When two or more types are present the effect of each amount and type separately is additive. Fig. 1b is then entered giving R', the first estimate of the range. Using 6 a.m. dry bulb and 1700Z 2000 ft wind velocity. Figs. 1c and d are next entered giving corrections D_{1} and D_{2}. The forecast range is then computed from the expression:

\[ R = R' + D_{1} + D_{2}. \]

The relative frequency distribution of errors in this dependant data is shown in Fig. 2 along with other relevant parameters of the distribution.

![Graph showing relative frequency distribution of errors.](image)

**Fig 2.** Relative frequency distribution of errors.
Note: The observed 1700Z 2000 ft wind was not in all cases used in the calculation of a forecast maximum. In two cases out of the 74 a shallow high pressure cell or belt with south-easterlies at sea level over Adelaide was the predominant feature of the surface wind system. However at 2000 ft the wind changed abruptly, in the one case to a light northwest in advance of a trough, and the other to a light southerly in the rear of a similar trough. As the observed wind thus appeared inhomogeneous it was discarded and the surface isobaric gradient wind calculated and used.

This appears justifiable as the 2000 ft wind was chosen purely to obtain a consistent estimate of the wind, for the advective component of the temperature range.

A similar occurrence took place twice in February, 1958, so it is well when operating with the method to observe whether the 2000 ft pilot balloon wind is in reasonable agreement with the M.S.L. pressure gradient, which is generally the deciding wind stream for low level temperature advection.

4. CONCLUSION

With regard to maximum temperature forecasting for special purposes, from data available at 6 a.m., it appears that the objective procedure described can produce a forecast with a useful error variance on days of no major air mass change and no rain. For these days, in particular for the change day, an estimate can first be made on the maximum temperature possible and then a subjective reduction in the range made, taking into account the estimated time of arrival of the change. For rain days there is no known deviation from the method and a subjective estimate is probably the best alternative.

Extension to other months of the year appears simply routine as heat sources and other effects are expected to change systematically during the year. However, it is advisable to check on the governing variables, as there is some justification for a change in the main dependant variate, i.e. range, and the number of "independent" variates where the natural variance of maximum temperature decreases significantly, as it does during winter.

Midsummer was chosen for the test period since maximum temperatures have their greatest variability in Adelaide at this time of the year.

Another useful extension of the objective procedure is concerned with the issue of maximum temperature forecasts for the following day using base data at about 3 p.m. or 9 p.m. local time the day before. Here considerable success has been obtained by forecasting
conditions at 6 a.m. the following day and applying these to the prediction graphs. This was facilitated in this case by using the results of a paper yet to be published by L.G. Veitch (C.S.I.R.O. Department of Mathematical Statistics) on minimum temperature forecasting. Wind forecasts were made by subjectively tempering known three hourly isallobaric rates and applying them to the known pressure distribution to obtain a forecast of gradient wind at 1700Z (2.30 a.m. local time).

Cloud forecasts were usually made using a known local subjective relation of wind stream to low cloud and general synoptic movement for middle and high clouds after sketching lines of equal cloud amount and type.

These extra calculations need of course only be done in the region of weather influence for the station in question. It is held, however, that when forecasting maximum temperatures for the following day, it is better to attempt to correlate objectively directly maximum temperature today with expected maximum temperature tomorrow, than to add one objective forecast of 6 a.m. temperature and one of range to obtain an estimated maximum. This is open to doubt if both methods are objective and physically consistent, and can apparently only be resolved by comparison of results. Once governing parameters have been established the method is simple in application and it is logical to apply the method to forecasting minimum temperatures or to frost forecasting.

In fact this graphical approach to objective forecasting can be used in many local objective forecasting studies where one variable such as temperature, cloud, wind and probably dew point is required to be forecast quantitatively.

ACKNOWLEDGMENTS

Appreciation is expressed to the senior members of the staff of the Adelaide Weather Bureau for their interest and encouragement in the work as it progressed and for their constructive criticisms of the final article.

Thanks are also due to the Research Section of the Central Bureau of Meteorology for useful guidance and also to the Division of Mathematical Statistics, C.S.I.R.O. (Adelaide University) for calculation of the regression co-efficients from the correlation matrix.
<table>
<thead>
<tr>
<th>Name</th>
<th>Year</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flower, W.D.</td>
<td>1937</td>
<td>Geophysical Memoirs, No. 71, p. 66.</td>
</tr>
<tr>
<td>Neiburger, M.</td>
<td>1941</td>
<td>B.A.M.S. March, 1941, p. 95.</td>
</tr>
</tbody>
</table>